DOCUMENT RESUME

ED 090 974 IR 000 585

AUTHOR	Larsson, Bernt
TITLE	The Influence of Scale Transformations: A Study of Factor Analysis on Simulated Data.
INSTITUTION	School of Education, Malmo (Sweden). Dept. of Educational and Psychological Research.
PUB DATE	Feb 74
NOTE	18p.; Didakometry No. 40
EDRS PRICE	MF-$0.75 HC-$1.50 PLUS POSTAGE
DESCRIPTORS	*Factor Analysis; *Factor Structure; Measurement; *Multidimensional Scaling; Statistical Analysis; Statistical Data; *Transformations (Mathematics)

ABSTRACT

This report describes a test of the robustness of factor-analytic methods in the face of various types of scale transformations on the data. Because of the complexities that would be involved in an exact analytical investigation, the tests were done with simulated sets of data having different factor structures. After factor analyzing the original data sets, scale transformations were done, and the transformed data sets were factor analyzed. Comparisons made between results obtained before and after transformations lead to the conclusion that monotonic transformations do not alter the results, while nonmonotonic transformations may. Because the comparisons were made with only a small number of data sets, it is suggested that special choices of data, factor-analytic methods, or scale transformations may limit the validity of this conclusion. (WDR)

DEPARTMENT OF
EDUCATIONAL AND
PSYCHOLOGICAL RESEARCH
SCHOOL OF EDUCATION
MALMÖ, SWEDEN

didakometry

Larsson, B.

THE INFLUENCE OF SCALE TRANSFORMATIONS: A STUDY OF FACTOR ANALYSIS ON SIMULATED DATA

No. 40 February 1974

THE INFLUENCE OF SCALE TRANSFORMATIONS: A STUDY OF FACTOR ANALYSIS ON SIMULATED DATA

Bernt Larsson

Larsson, B. The influence of scale transformations: A study of factor analysis on simulated data. Didakometry (Malmö, Sweden: School of Education), No. 40, 1974.

Data with different factor structures are generated and analyzed. The variables are transformed and reanalyzed and comparisons between factor analyses before and after transformation are made. All comparisons indicate the same conclusion: monotonic transformations do not change the results, while non-monotonic transformations may . Special choices of data, factor-analytic method, transformations and ways of comparison may limit the validity of this conclusion.

Keywords: Measurements, transformations, factor analysis

INTRODUCTION

Educational research uses many concepts which are not unequivocally defined. This has involved different variables to measure the (nominally) same property. Relations between these are almost always stochastic, ranging from complete independence to the maximal correlation "allowed" by their reliabilities. Since several variables, proposed to measure the same property, are seldom congeneric and often not even isomorphic in their specific true scores, they can hardly be said to measure the same property.

However, from a practical viewpoint, it need not be important that variables measure exactly the same property. It is more important that they represent the same property to a sufficient extent. By this I mean that the same result is obtained by different collections of variables, which are considered to measure the same properties, when they are used on the same or similar measurement objects. This is a vaguely formulated but important principle. Above all it means that a researcher draws the same conclusions, generates the same hypotheses and makes the same decisions, independent of which collection of variables the results are based on. At the present state of educational measurements it is, no doubt, of importance to investigate the robustness of results based on different collections of variables. Such studies can never be definitive, but this report tries to give some results relevant to the question of robustness.

These investigations can be performed in different ways: one can use real data or simulated data or one can make a purely analytical (mathematical) investigation. Real data have the advantage of permitting concrete interpretations. The drawback is that you, as a rule, must take data already collected, which often are designed for quite another purpose. It is considerably simpler to make a systematic investigation by simulating data, since data can be chosen almost without restriction. However, the amount of data, which can be reasonably analyzed, limits the possibility of generalizing from simulation experiments. An analytical investigation is here superior, since it is not based upon special data. Owing to complex problems, analytical investigations are not always feasible.

I have chosen to study the influence of some transformations on factor-analytical results. As an example, suppose that results of factor analysis are robust to monotonic transformations. It would then seem to me, that the scale problem of the instruments chosen is not very urgent, provided that the properties are defined sufficiently well to determine ordinal measurements. Strictly speaking, this study investigates only the robustness of allocating different numbers to the possible outcomes of a certain collection of instruments. But since functions can approximate stochastic relations this report also mirrors, more or less, the robustness of results based on different collections of instruments.

The present report comprises simulations only. In my opinion, it would have been better to make an analytical investigation. However, the complexity of the problems is clearly too great for me - I do not even know how they should be formulated. Simulation is a solution which can be resorted to when an analytical investigation does not seem possible. The results of this report therefore constitute no rigid proof either for or against factor-analytical robustness to transformations: they can only make it more or less credible.

DESIGN OF THE EXPERIMENT

Thurstone (1947, p. 369) says that comparisons have shown that different monotonic transformations give essentially the same factor structure, when this is a simple structure. However, he does not show any results. His statement is, in a sense, corroborated by his box example, which can be found in several places in his book. If one knows that a collection of variables satisfies the linear factor model, then monotonic transformations of these variables cannot satisfy the model in the same way. The variables of the box example are different measurements of boxes, which often are non-linear functions of height, length and breadth. In spite of this, factor loadings of the rotated factors give support for the above-mentioned dimensions. Thus, there are certain reasons to assume the robustness of factor analysis, at least to monotonic transformations.

As a further support for the same presumption, one may add the following simple, analytic result. Suppose that y_1 and y_2 have a bivariate normal distribution with expected values μ_1 and μ_2, variances σ_1^2 and σ_2^2 and correlation ϱ. Then

$$(1)\quad \varrho_{y_1^2 y_2} = \pm \varrho \sqrt{\frac{2}{\vartheta_1^2 + 2}}$$

and

$$(2)\quad \varrho_{y_1^2 y_2^2} = \frac{\varrho(\varrho \vartheta_1 \vartheta_2 + 2)}{\sqrt{(\vartheta_1^2 + 2)(\vartheta_2^2 + 2)}}$$

Here $\vartheta = \sigma/\mu$, the coefficient of variation, and $\varrho_{y_1 y_2^2}$ shall have the same sign as $\mu_1 \varrho$. As ϑ usually lies in the interval (0.0, 0.5) for variables within behavioural research, a quadratic transformation has little effect on ϱ. One dare assume that such a transformation hardly changes results of factor analysis on approximately normally distributed variables. It should be pointed out that y^2 is here not a strictly monotonic transformation, since $-\infty \leq y \leq \infty$. The extreme case $z^2 = [(y - \mu)/\sigma]^2$ is clearly non-monotonic and formulas 1 and 2 show that $\varrho_{z_1^2 y_2} = 0$, independent of ϱ, and $\varrho_{z_1^2 z_2^2} = \varrho^2$. This indicates that non-monotonic transformations drastically can change the factor structure.

A number of collections of variables with known factor structures have been generated. Then these have been transformed in different ways and factor analyses before and after transformation are compared in some aspects. The original variables have been generated by the following factor model:

$$(3) \qquad y_i = \sum_{k=1}^{m} b_{ik} x_k + e_i \quad , \; i = 1, \ldots, p.$$

Here y_i is a manifest variable, x_k a common factor, e_i a unique factor and b_{ik} a factor loading. The model may now be realized in various ways. I have chosen to make x_k and e_i ($p + m$ in number) independent of each other. Also, all x_k are so called stanine variables (approximatively normally distributed on the integers 1(1)9), all e_i are rectangularly distributed on the interval (0, 2) and $b_{ik} \geq 0$. This gives all $y_i \geq 0$, symmetrically distributed with negative kurtosises, a distribution rather common within behavioural sciences. The linear correlation between y_i and y_j is now

$$(4) \qquad \varrho_{y_i y_j} = \frac{3.84 \sum_k b_{ik} b_{jk}}{\sqrt{(3.84 \sum_k b_{ik}^2 + 0.33)(3.84 \sum_k b_{jk}^2 + 0.33)}}$$

A ϱ value aimed at can be obtained by suitable choices of the factor loadings. As $b_{ik} \geq 0$, ϱ becomes positive, and this is a general fact for several variable domains. The choice of identical e_i implies the rank correlation between communality and variance of y_i to be unity. However, I do not think that this restriction makes the results less general.

For every collection of variables there remains, among other things, the choice of m, p and ϱ. As the number of factor loadings, which must be determined for a given collection, is mp, I have put some restrictions on the choice of p because of the amount of work involved: it is fixed to 10 and 30.

Ten manifest variables constitute a small number as far as factor analyses used in behavioural research are concerned, while thirty is a more common, though not especially large number. Then m has been chosen to indicate either a small or large number of factors: for $p = 10$, $m = 1$ or 4 and for $p = 30$, $m = 4$ or 10. For these four cases I have chosen correlation matrices in three ways: $0.0 \leq \varrho \leq 0.4$, $0.0 \leq \varrho \leq 0.9$ and $0.5 \leq \varrho \leq 0.9$.

The factor-analytic method used here is the principal axes solution with varimax rotation according to the program BMDX72, Dixon (1970). I think that this method is too much used. Whether this depends on tradition (most earlier analyses built upon the centroid method with graphical rotations to simple structures, of which the method of BMDX72 is a modern variant), easily available programs or difficulties in understanding newer methods may be left an open question. As is clear from e.g. Jöreskog's papers the newer, inferential methods of factor analysis are more stringent and flexible than the older ones and will, in my opinion, dominate future uses of factor analysis. Today they are more or less limited because of computers having insufficient internal memories. For instance, the maximal number of variables which can be analyzed is not seldom too small for applications within behavioural science. Inertia of innovation may be added to this: the inferential factor analysis puts some new demands on the user.

Although I could have used programs like ACOVS or LISREL, see e.g. Jöreskog (1973) and Jöreskog & van Thillo (1973), BMDX72 has been used, for two reasons. Partly because it is easily available but, above all, because so many researchers have used it (or more correctly: its parallel BMD03M). I do not think that results would have been essentially different, as far as the robustness of the estimates of parameters of the factor model is concerned, with a method other than that of BMDX72.

The program has been run twice for every case, partly with 1.0 and partly with squared multiple correlations (R^2) in the principal diagonal of the correlation matrix. There are a total of 24 factor analyses on untransformed (original) variables, which are designed as a 2x2x3x2 factorial experiment. However, there are only 12 different collections of variables generated, since the two types of values in the principal diagonal are used for the same generation. Table 1 shows the design with the numbering of cases which will be used when reporting the results.

Table 1. Numbering of the different cases

		p = 10		p = 30	
		m = 1	m = 4	m = 4	m = 10
$0.0 < \varrho < 0.4$	R^2	1A	4A	7A	10A
	1.0	1B	4B	7B	10B
$0.0 < \varrho < 0.9$	R^2	2A	5A	8A	11A
	1.0	2B	5B	8B	11B
$0.5 < \varrho < 0.9$	R^2	3A	6A	9A	12A
	1.0	3B	6B	9B	12B

Formula 4 is exactly valid when we have an infinite number of measurement objects and will be approximative when only a small number of objects is available. For instance, the factor variance is not exactly 3.84 and the factors are not exactly independent of each other. This implies that you cannot, in practice, obtain exactly the ϱ values aimed at, e.g. $\varrho = 0.0$ may very well be realized as -0.1. A compromise has to be made between reasonable costs for computer time and a sufficient number of objects to approximate the model. Trial runs with 50, 100 and 200 objects showed that only 200 objects give acceptable agreements between model and data. Each of the 24 cases shown in table 1 is thus based on 200 measurement objects, a rather common sample size within educational research. The cases have been generated twice in order to get an idea of random variation at 200 objects. Information about this variation will be used when reporting the result.

The number of possible transformations is infinite. With regard to computer time and the amount of work when comparing factor analyses, the number must be strongly limited. My choice is hardly very rational or systematic: I do not even know how it could be made so. Positively skewed distributions are not unusual and it is sometimes recommended that these should be normalized through square root or logarithmic transformations. These functions have been exploited, as well as their inverse functions. As an example of a more general monotonic transformation, I have used rank numbers instead of the original scores. Finally, one non-monotonic transformation has also been chosen. Such transformations are sometimes used, e.g. absolute deviation from an ideal point on a scale.

The 12 untransformed cases (24 factor analyses) have given rise to four new transformed sets. For one set, half of the variables (those with odd numbers) have been transformed from y into $y^2/10$, while the other variables have been transformed from y into $\sqrt{y}$. For a second set, the corresponding functions are exp (y/6) and ln (1+y) and in a third set all variables have been ranked. The set involving non-monotonic transformations uses $z_y^2 = [(y - m_y)/s_y]^2$ for half of the variables of a case and leaves the other half unchanged. Every set comprises 24 factor analyses as every case is run twice, both with R^2 and 1.0 in the principal diagonal of the correlation matrix. Thus there are totally 144 factor analyses (the untransformed set is generated twice).

The way of comparing results of factor analyses is not self-evident. What is meant by saying that two analyses give essentially the same answer? Which aspects are to be compared and how? An important interpretation of "the same answer" is that different researchers understand the factors in a similar way. This configurational invariance is in most cases sufficient. The drawback of simulated data is the impossibility of empirically interpreting factors: data are, so to say, without content. One then has to examine numerical invariance by calculating different indices for deviation. This is a more rigorous comparison: e.g. numerical invariance of factor loadings implies configurational invariance but the reverse need not be true.

Comparisons will be concentrated on eigenvalues: the number of factors with eigenvalues above 1.0 (a common criterion used when rotating factors), the proportion of total variance accounted for by these factors and, above all, the distribution of eigenvalues of unrotated factors. Comparisons of communalities will also be commented upon, while factor loadings are discussed rather little. The other comparisons should still give the reader an understanding of the influence of the transformations.

RESULTS

The numbering of cases which was shown in table 1 is used in the following tables. The six sets of cases will be numbered by Roman numerals: I and II stand for the untransformed sets, III concerns the transformations $y^2/10$ and $\sqrt{y}$, IV refers to exp (y/6) and ln (1+y), V comprises the rank numbers and VI denotes the set with non-monotonic transformations. Set I has been used when generating the transformed sets. Comparisons between set I and sets III, IV, V and VI are therefore the most important ones. However, comparisons between I and II give you a hint of the size of random variations and can be exploited for discussions of the other comparisons.

The first aspect of comparison concerns the number of unrotated factors with eigenvalues greater than 1.0. This is a "blind" criterion which is often, perhaps too often, used when determining the number of factors to rotate. If you have no prior idea about the number of interpretable factors, it is perhaps wiser to examine more than one solution. It is far from certain that a factor with an eigenvalue of 1.5 can be given any interpretation, while I have sometimes seen how a factor with an eigenvalue below 1.0 has contributed essentially to the understanding of a variable domain. Dempster (1969, p. 139) has a similar, though more extreme attitude concering component analysis.

Table 2 gives the numbers of unrotated factors with eigenvalues exceeding 1.0 and table 3 shows the proportion of total variance which these factors represent. It is clear from these tables that III, IV and V are in very good agreement with I, and the difference between I and II is also small. As might have been expected, VI shows greater deviation. The number of factors to rotate is, with one exception, equal to or greater than that for I, but in spite of this fact the proportion is often lesser for VI than for I. Thus, set VI has a flatter eigenvalue distribution than I, which is reasonable with regard to formulas 1 and 2. However, not even as extreme a transformation as VI comprises can be said to produce very great differences. But tables 2 and 3 present very rough measures: they tell rather little about similarities or differences between corresponding factors of two sets.

Table 2. Number of factors with eigenvalues above 1.0

Case	Set I	II	III	IV	V	VI
1A	1	1	1	1	1	1
1B	3	3	3	3	3	4
2A	1	1	1	1	1	2
2B	1	2	1	1	1	3
3A	1	1	1	1	1	2
3B	1	1	1	1	1	2
4A	1	1	1	1	1	1
4B	3	4	3	3	3	4
5A	1	1	1	1	1	2
5B	2	2	2	2	2	3
6A	1	1	1	1	1	2
6B	1	1	1	1	1	2
7A	2	2	2	2	2	2
7B	9	10	10	9	10	13
8A	2	2	2	2	2	2
8B	5	5	5	5	5	7
9A	2	2	2	2	2	2
9B	2	3	2	2	2	4
10A	2	2	2	2	2	2
10B	11	10	10	11	11	11
11A	3	4	3	3	3	2
11B	5	5	5	5	6	7
12A	3	5	3	3	3	4
12B	5	5	5	5	5	6

Table 3. Proportion of total variance for factors with eigenvalues above 1.0

Case	Set					
	I	II	III	IV	V	VI
1A	0.222	0.162	0.213	0.214	0.210	0.124
1B	0.515	0.471	0.507	0.508	0.505	0.553
2A	0.610	0.618	0.593	0.594	0.598	0.501
2B	0.638	0.742	0.622	0.625	0.626	0.690
3A	0.744	0.754	0.716	0.722	0.721	0.640
3B	0.768	0.777	0.742	0.748	0.745	0.710
4A	0.210	0.205	0.209	0.209	0.194	0.104
4B	0.504	0.607	0.502	0.503	0.489	0.544
5A	0.538	0.537	0.534	0.526	0.522	0.424
5B	0.679	0.681	0.677	0.670	0.664	0.630
6A	0.686	0.674	0.677	0.672	0.662	0.572
6B	0.711	0.700	0.704	0.700	0.690	0.659
7A	0.211	0.230	0.209	0.211	0.192	0.139
7B	0.633	0.606	0.590	0.559	0.639	0.643
8A	0.586	0.592	0.579	0.575	0.564	0.465
8B	0.727	0.731	0.722	0.718	0.708	0.687
9A	0.741	0.712	0.734	0.727	0.729	0.606
9B	0.753	0.762	0.746	0.740	0.741	0.703
10A	0.212	0.235	0.210	0.211	0.195	0.129
10B	0.622	0.604	0.587	0.622	0.609	0.568
11A	0.567	0.595	0.561	0.555	0.533	0.429
11B	0.674	0.672	0.669	0.664	0.678	0.662
12A	0.704	0.764	0.699	0.689	0.692	0.549
12B	0.799	0.797	0.795	0.787	0.787	0.681

Only the first five eigenvalues of unrotated factors have been exploited for comparisons of eigenvalue distributions. The deviations of the subsequent eigenvalue pairs are small throughout: the greatest deviation almost always belongs to the first eigenvalue pair. The sum of the absolute differences of the first five eigenvalue pairs is presented as an index of deviation. The sum is then, of course, an upper limit for individual differences and it has been calculated for differences between eigenvalues of set I and those in II, III, IV, V and VI. These sums are given in table 4 with certain summaries in table 5.

The transformations of III and IV almost always cause small deviations, smaller than those of a new data generation (set II). Set V involves deviations of the same magnitudes as for II, sometimes somewhat smaller and sometimes a bit greater. The non-monotonic transformation of VI on the other hand gives rise to great deviations. These depend mainly on the fact that the first factor for every case of I has great loadings on most variables, while, for VI, the first factor has only great loadings on y variables and the second factor is defined by the z_y^2 variables. Those variables, which have been transformed from y to z_y^2, consequently measure something else now.

Table 4. Eigenvalue deviation from set I

Case	Set				
	II	III	IV	V	VI
1A	0.724	0.145	0.115	0.163	1.188
1B	0.701	0.138	0.094	0.179	1.439
2A	0.174	0.244	0.218	0.249	4.422
2B	0.246	0.222	0.163	0.231	4.899
3A	0.118	0.392	0.318	0.303	5.969
3B	0.156	0.447	0.326	0.361	6.527
4A	0.222	0.043	0.024	0.202	1.299
4B	0.160	0.026	0.012	0.194	1.211
5A	0.102	0.070	0.169	0.245	3.970
5B	0.158	0.034	0.112	0.235	4.383
6A	0.170	0.146	0.253	0.292	5.088
6B	0.257	0.119	0.176	0.351	5.873
7A	0.637	0.101	0.087	0.626	2.664
7B	0.587	0.083	0.075	0.638	2.802
8A	0.653	0.243	0.367	0.722	12.620
8B	0.309	0.252	0.345	0.724	12.504
9A	1.316	0.271	0.598	0.509	16.464
9B	1.464	0.260	0.476	0.499	16.997
10A	1.322	0.068	0.031	0.514	2.550
10B	1.338	0.080	0.031	0.530	2.658
11A	0.954	0.208	0.429	1.167	10.223
11B	0.741	0.169	0.306	1.219	10.239
12A	1.279	0.151	0.456	0.482	11.009
12B	1.261	0.146	0.456	0.481	11.760

According to table 5 the average deviation, for a given set, is the same irrespective of whether R^2 or 1.0 has been used. There is an indication of greater random error with more variables (comparison I-II) and that VI and perhaps also V deviate more for p = 30 than for p = 10. It seems to have no importance whether the number of factors is small or great. For III, IV and VI, transformations have a greater influence on high than on low correlations, which seems reasonable considering formulas 1 and 2. However, we must not forget that the transformations of III and IV have almost no influence: factor analysis on y or e.g. y^2 gives in principle the same result.

Table 5. Summary of table 4

Summary		Set				
		II	III	IV	V	VI
Diagonal value	R^2	0.639	0.173	0.255	0.456	6.455
	1.0	0.615	0.165	0.213	0.470	6.774
p	10	0.266	0.169	0.165	0.250	3.856
	30	0.988	0.169	0.303	0.676	9.374
m	low	0.590	0.233	0.265	0.434	7.375
	high	0.664	0.105	0.203	0.492	5.855
ε	low	0.711	0.086	0.059	0.381	1.976
	medium	0.418	0.180	0.264	0.599	7.908
	high	0.752	0.242	0.380	0.409	9.961
Total		0.627	0.169	0.234	0.463	6.615

And now some words about the squared multiple correlation between y_i and $x_1 \ldots x_m$, the so called communality (Γ_i). It is known (see e.g. Rozeboom, 1966, p. 261) that the squared multiple correlation between y_i and $y_1, \ldots, y_{i-1}, y_{i+1}, \ldots, y_p$ (P_i^2) cannot exceed Γ_i, when these quantities are based on an infinite number of objects. The relation may be another in a sample and I have examined whether the sample value of P_i^2, which is an estimate of Γ_i often used, seems to be good for this purpose. The communality is known for I and II and I have looked through set I with the result that R_i^2 presumably can be said to constitute a reasonable estimate of Γ_i. Few differences are greater than 0.10 and $R_i^2 - \Gamma_i$ is greatest for low R_i^2 values, because the bias $R_i^2 - P_i^2$ is then not negligible.

A simple investigation of whether the transformations change Γ_i has also been made. But Γ_i is not known for the transformed sets and comparisons have therefore been made on sample values calculated from the factor loading matrix. (The matrix has m factors, except for the cases with m = 10, where only five factors have been used.) The same pattern as for eigenvalues comes back. The monotonic transformations have hardly any influence but the non-monotonic does. Deviations over 0.05 are rare for III, IV and V, while deviations of 0.20 are not unusual for VI: maximal deviation per case varies here between 0.22 and 0.84.

Though eigenvalues and communalities do not deviate from each other (comparisons I-III, I-IV and I-V) this does not usually imply that factor loadings must be similar too. However, a superficial inspection of these (for unrotated factors) shows no new picture. Factor loadings of set I are for every case similar to corresponding loadings of sets III, IV and V: a difference over 0.10 is a rarity. On the other hand, loadings are differently structured in VI and great differences are common. There is therefore reason to presume that factors of I, III, IV and V - but not of VI - would have been interpreted in similar ways if the variables had had some empirical anchoring.

DISCUSSION

As I see it, a measurement process consists of three stages: a definition of a property, a choice of an instrument and the allocation of numbers to the possible outcomes of the instrument. In some areas researchers have been able to agree upon a definition of a property so precise that all admissible combinations of instruments and numbers determine linearly related variables. This is hardly the case in educational research. I believe that the definitions are here sometimes so diffuse that possible combinations made by researchers, believing in the same definition, generate variables, which are not even monotonically related in their specific true scores. That is, the difference in one true score variable has not the same sign as the difference in another true score variable for any pair of objects. I have sometimes heard pronouncements like "this is probably only an ordinal scale" just as if it should be self-evident that a variable represents a property according to the requirements of the ordinal scale. Strictly speaking, the pronouncement is a contradiction e.g. every time a researcher constructs two alternative instruments for measuring a property and does not find the two true score variables monotonically related, in the sense that both instruments do not produce ordinal scales for the same property.

However, it may well also be so that stochastic relations, which essentially approximate monotonic relations (e.g. 90 per cent of pairs of objects have differences with the same sign for two variables), are sufficient for several analytical purposes But this robustness of results based on different collections of variables is something that we know rather little about. What we would like to know is in which situations robustness occurs - and does not occur. This information could then be used to focus our efforts to improve measurements for situations where robustness does not exist. We may imagine a two-dimensional contingency table with different properties as columns and different statistical methods as rows and where every cell can be said to define a situation. My presumption is that some methods are more robust than others, e.g. a linear product-moment correlation seems to be much more scale independent than a statistical test about equal covariance matrices. Likewise, properties may also vary in robustness: diffusely defined properties generate less robustness because the admissible choices of variables are so great. This would mean that some properties may be sufficiently well defined for certain methods but not for others. A wise selection of methods and properties could form a basis for a research program of empirical investigations of robustness.

My simulation experiment is not tied to certain properties but to one method and thus more or less mirrors the conditions for a whole row of the above-mentioned contingency table. In that it is presumably more general than an empirical investigation. On the other hand, the experiment involves the restriction of only examining non-stochastic relations, which means that it primarily treats the robustness of the third stage of the measurement process: the allocation of numbers, given a certain collection of instruments. The special choices of factor-analytic method, ways of comparison and transformations may also restrict the generalizability of this investigation. I will briefly comment upon these choices.

The question as to whether another method of factor analysis would have produced different results seems to be difficult to answer. I would like to answer in the negative, but this is only what I believe. In my experience, several descriptive methods seem to be rather robust to many (but not necessarily all) monotonic transformations, while inferential methods need not be. More exactly: several estimates are often little dependent on the form of the distribution, but the probabilistic evaluation of a statistical test quantity can be very sensitive to different kinds of distribution functions. I therefore believe that the results obtained in this report would not have been essentially different if the estimates had been produced by another technique, say maximum likelihood factor analysis.

Man has a limited conception of multidimensional phenomena, at least if they involve more than three dimensions. One may argue that it is rather meaningless to present descriptions more or less void of characteristics that can be exploited, even if they are of interest to statistical theory. (For instance, I find the determinant of a covariance matrix much less understandable than its trace.) It is not easy to see what kinds of comparisons will be the most fruitful ones to undertake in factor analysis, especially since simulated data have no empirical anchoring. I have chosen to focus the comparisons on some characteristics commonly used in factor analysis. The evaluation of the comparisons between untransformed and transformed data has also been facilitated by a second generation of data (set II). Since all the comparisons made seem to tell the same story there are reasons to believe that other numerical comparisons would not have altered the results. But it would be valuable to supplement this report with parallel investigations on real data in order to elucidate the influence of transformations on the interpretation of factors.

Larsson (1973) gives an example where a correlation has been considerably changed by monotonic transformations. Similar results have been obtained by others, e.g. Box & Cox (1964) or Kruskal (1965). It is not easy to state under what conditions correlations can be changed much or little by monotonic transformations, but I suppose that there exist correlation matrices containin several correlations which can be changed appreciably so that the factor structure will also change. I have no idea at all whether this happens frequently or not. We must not forget that I have chosen some rather common monotonic transformations independent of data. They are certainly not optimal in the sense that they change the factor structure as much as possible On the other hand, it may well be so that in many cases the maximal change is negligible. (Notice that the robustness of data to the rank transformation does not imply robustness to any monotonic transformation. The rank transformation is dependent on the distribution, e.g. a rectangular distribution implies no change at all.) Therefore it is perhaps wisest to state a conditional conclusion: the monotonic transformations used show hardly any non-robustness of factor analysis.

If the results obtained in this study should occur often, methods like nonmetric and nonlinear factor analyses (see e.g. Lingoes & Guttman, 1967, Carroll, 1972, and McDonald, 1962) would seldom be necessary. This is in line with what Shepard (1972, p. 37) says about his and Kruskal's nonmetric variety of factor analysis: " ... it has never been widely used. ... the method tends-except in the case of extremely nonlinear data - to yield representations that differ but little from those obtained by classical (linear) factor analysis."

REFERENCES

Box, G.E.P. & Cox, D.R. An analysis of transformations. J. roy. statist. Soc. B, 1964, 26, 211-252.

Carroll, J.D. Individual differences and multidimensional scaling. In: Shepard, R.N., Romney, A.K. & Nerlove, S.B. (Eds.) Multidimensional scaling. Theory and applications in the behavioral sciences. Vol. I. New York: Seminar Pr., 1972. Pp. 105-155.

Dempster, A.P. Elements of continuous multivariate analysis. Reading, Mass.: Addison-Wesley, 1969.

Dixon, W.J. (Ed.) BMD. Biomedical computer programs. X-series supplement. Berkeley, Calif.: Univer. Calif. Pr., 1970.

Jöreskog, K.G. Analyzing psychological data by structural analysis of covariance matrices. Research Report 73 - 2. Department of Statistics, University of Uppsala, 1973.

Jöreskog, K.G. & van Thillo, M. LISREL. A general computer program for estimating a linear structural equation system involving multiple indicators of unmeasured variables. Research Report 73 - 5. Department of Statistics, University of Uppsala, 1973.

Kruskal, J.B. Analysis of factorial experiments by estimating monotone transformations of the data. J. roy. statist. Soc. B, 1965, 27, 251-263.

Larsson, B. Obtaining maximal correlations by the construction of binary variables. Didakometry, No. 38, 1973.

Lingoes, J.C. & Guttman, L. Nonmetric factor analysis: a rank reducing alternative to linear factor analysis. Multiv. Behav. Res., 1967, 2, 485-505.

McDonald, R.P. A general approach to nonlinear factor analysis. Psychometrika, 1962, 27, 397-415.

Rozeboom, W.W. Foundations of the theory of prediction. Homewood, Ill.: Dorsey Pr., 1966.

Shepard, R.N. A taxonomy of some principle types of data and of multidimensional methods for their analysis. In: Shepard, R.N., Romney, A.K. & Nerlove, S.B. (Eds.) Multidimensional scaling. Theory and applications in the behavioral sciences. Vol. I. New York: Seminar Pr., 1972. Pp. 21-47.

Thurstone, L.L. Multiple-factor analysis. Chicago: Univer. Chicago Pr., 1947.

Department of
Educational and
Psychological Research
School of Education
Malmö, Sweden

Abstract card

Larsson, B. The influence of scale transformations: A study of factor analysis on simulated data. _Didakometry_ (Malmö, Sweden: School of Education), No. 40, 1974.

Data with different factor structures are generated and comparisons between factor analyses before and after transformations are made. All comparisons indicate the same conclusion: monotonic transformations do not change the results, while non-monotonic transformations may do.

Indexed:
1. Measurements
2. Transformations
3. Factor analysis

Reference card

Larsson, B. The influence of scale transformations: A study of factor analysis on simulated data. _Didakometry_ (Malmö, Sweden: School of Education), No. 40, 1974.